

Generating Videos with Dynamics-aware Implicit Generative Adversarial Networks

Year: 2022 | Citations: 216 | Authors: Sihyun Yu, Jihoon Tack, Sangwoo Mo, Hyunsu Kim, Junho Kim

Abstract

In the deep learning era, long video generation of high-quality still remains challenging due to the spatio-temporal complexity and continuity of videos. Existing prior works have attempted to model video distribution by representing videos as 3D grids of RGB values, which impedes the scale of generated videos and neglects continuous dynamics. In this paper, we found that the recent emerging paradigm of implicit neural representations (INRs) that encodes a continuous signal into a parameterized neural network effectively mitigates the issue. By utilizing INRs of video, we propose dynamics-aware implicit generative adversarial network (DIGAN), a novel generative adversarial network for video generation. Specifically, we introduce (a) an INR-based video generator that improves the motion dynamics by manipulating the space and time coordinates differently and (b) a motion discriminator that efficiently identifies the unnatural motions without observing the entire long frame sequences. We demonstrate the superiority of DIGAN under various datasets, along with multiple intriguing properties, e.g., long video synthesis, video extrapolation, and non-autoregressive video generation. For example, DIGAN improves the previous state-of-the-art FVD score on UCF-101 by 30.7% and can be trained on 128 frame videos of 128x128 resolution, 80 frames longer than the 48 frames of the previous state-of-the-art method.